

# Thirty Meter Telescope (TMT) Narrow Field Infrared Adaptive Optics System (NFIRAOS) real-time controller preliminary architecture

Dan Kerley<sup>a</sup>, Malcolm Smith<sup>a</sup>, Jennifer Dunn<sup>a</sup>, Glen Herriot<sup>a</sup>, Jean-Pierre Véran<sup>a</sup>, Corinne Boyer<sup>b</sup>, Brent Ellerbroek<sup>b</sup>, Luc Gilles<sup>b</sup>, and Lianqi Wang<sup>b</sup>

<sup>a</sup>National Research Council of Canada - Herzberg, 5071 West Saanich Rd, Victoria, Canada

<sup>b</sup>Thirty Meter Telescope, Suite 300 - 100 West Walnut Street, Pasadena, United State of America

## ABSTRACT

The Narrow Field Infrared Adaptive Optics System (NFIRAOS) is the first light Adaptive Optics (AO) system for the Thirty Meter Telescope (TMT). A critical component of NFIRAOS is the Real-Time Controller (RTC) subsystem which provides real-time wavefront correction by processing wavefront information to compute Deformable Mirror (DM) and Tip/Tilt Stage (TTS) commands. The National Research Council of Canada - Herzberg (NRC-H), in conjunction with TMT, has developed a preliminary design for the NFIRAOS RTC.

The preliminary architecture for the RTC is comprised of several Linux-based servers. These servers are assigned various roles including: the High-Order Processing (HOP) servers, the Wavefront Corrector Controller (WCC) server, the Telemetry Engineering Display (TED) server, the Persistent Telemetry Storage (PTS) server, and additional testing and spare servers.

There are up to six HOP servers that accept high-order wavefront pixels, and perform parallelized pixel processing and wavefront reconstruction to produce wavefront corrector error vectors. The WCC server performs low-order mode processing, and synchronizes and aggregates the high-order wavefront corrector error vectors from the HOP servers to generate wavefront corrector commands. The Telemetry Engineering Display (TED) server is the RTC interface to TMT and other subsystems. The TED server receives all external commands and dispatches them to the rest of the RTC servers and is responsible for aggregating several offloading and telemetry values that are reported to other subsystems within NFIRAOS and TMT. The TED server also provides the engineering GUIs and real-time displays. The Persistent Telemetry Storage (PTS) server contains fault tolerant data storage that receives and stores telemetry data, including data for Point-Spread Function Reconstruction (PSFR).

**Keywords:** TMT, NFIRAOS, RTC, MCAO, AO, real-time, adaptive optics

## 1. INTRODUCTION

The Narrow Field Infrared Adaptive Optics System (NFIRAOS) is a Laser Guide Star (LGS) Multi-Conjugate Adaptive Optics (MCAO) system for the Thirty Meter Telescope (TMT) that provides correction of atmospheric and telescope image aberrations over a 2 arc-minute field of view in the near infrared.<sup>1,2</sup> It will be located on the Nasmyth platform of TMT and will feed light to any one of three selectable client instruments. NFIRAOS performs wavefront correction using two deformable mirrors (DM) conjugated to 0km and 11.8km, with one DM mounted on a tip/tilt stage (TTS). Within NFIRAOS are six LGS Shack-Hartmann Wavefront Sensors (WFS) and one high-order natural guide star Pyramid Wavefront Sensor (PWFS). Additional wavefront information is provided to NFIRAOS by up to three Shack-Hartmann On-Instrument Wavefront Sensors (OIWFS) located in the science instruments, and up to four On-Detector Guide Windows (ODGW) on the science imager.

The control software for NFIRAOS is partitioned into two major subsystems: the NFIRAOS Real-Time Controller (RTC) and the NFIRAOS Components Controller (NCC). The NCC is responsible controlling all slower

---

Further author information: (Send correspondence to Dan Kerley)

Dan Kerley: E-mail: dan.kerley@nrc-cnrc.gc.ca, Telephone: 1-250-363-8107

NFIRAOS mechanisms and other non-real-time operations. Conversely, the RTC is responsible for providing real-time wavefront correction by processing all wavefront information to compute DM and TTS commands, and providing additional data streams to other subsystems, including: the Telescope Control System (TCS); the Laser Guide Star Facility (LGSF); Point-Spread Function Reconstructor (PSFR); client instruments; and the NCC, as shown in Figure 1. The RTC also provides telemetry data to the Reconstructor Parameter Generator (RPG) subsystem, which in turn processes the data to generate new control parameters for the RTC, including: reconstruction matrices; signal gains; filter parameters; and other values. Both the NCC and RTC are sequenced, at a high level, by the TMT Adaptive Optics Sequencer (AOSQ), which is responsible for configuring and sequencing all aspects of TMT AO systems.<sup>3</sup>

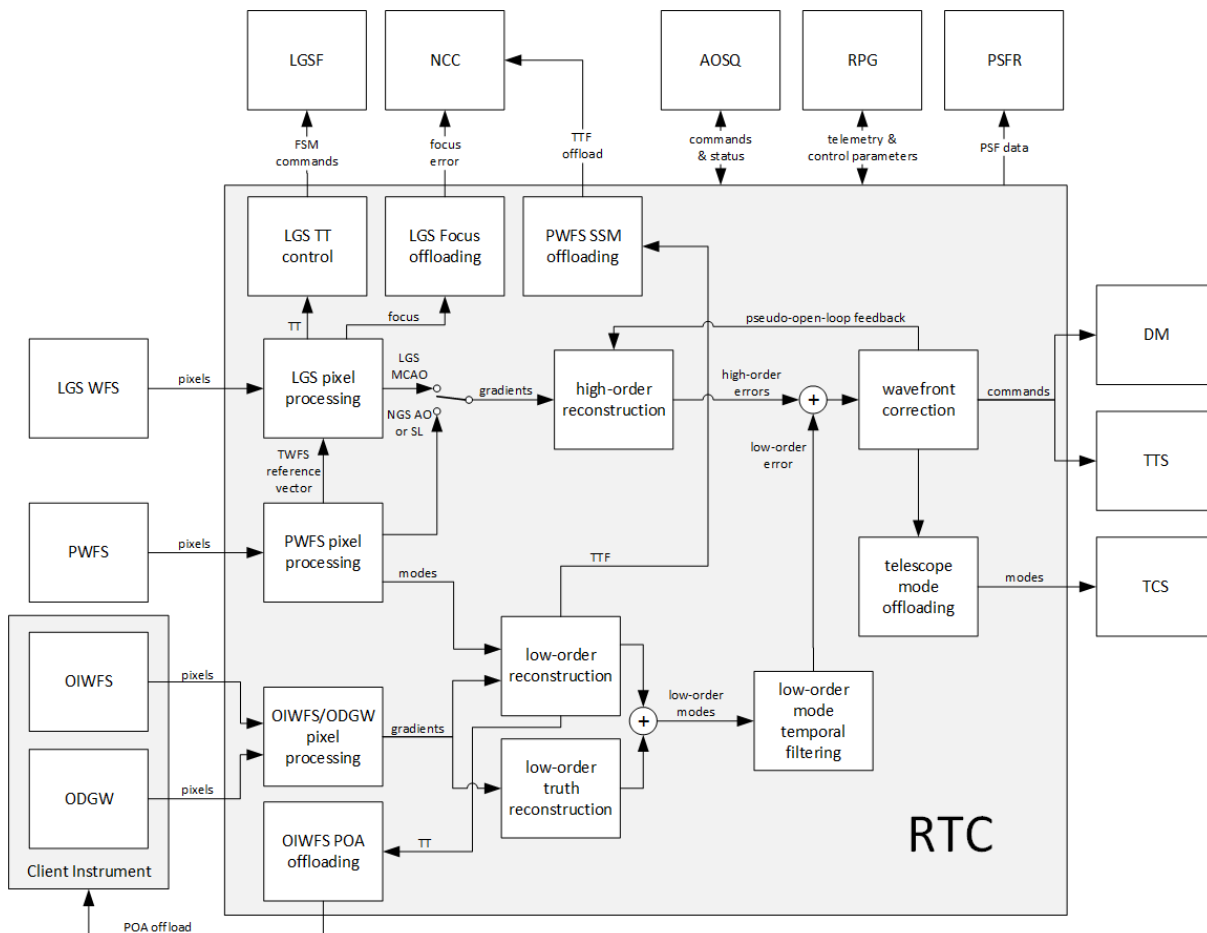


Figure 1. The primary purpose of the RTC is to process wavefront information and compute DM and TTS commands, however the RTC is also responsible for several offload tasks. Not all data flows are shown in this simplified block diagram.

The preliminary architecture for the RTC is comprised of nine Linux-based servers.<sup>\*</sup> Up to six servers are assigned the High-Order Processing (HOP) role, while one server each will be assigned the role of Wavefront Corrector Controller (WCC), the Telemetry Engineering Display (TED), the Persistent Telemetry Storage (PTS). All RTC servers, excluding the PTS, will be located on the Nasmyth platform in the NFIRAOS electronics enclosure. The PTS server will be located in the TMT computer room.

<sup>\*</sup>The design presented in this paper has been slightly updated since the preliminary design review and reflects the current thinking with respect to the system architecture and hardware selection.

## 2. SOFTWARE ARCHITECTURE

While the RTC can be thought of as single unified system, it is in fact a collection of nine servers, not including spares and test equipment, working in parallel to achieve the desired performance. Each of these servers has one of the four unique roles within the RTC:

- High-Order Processing (HOP)
- Wavefront Corrector Controller (WCC)
- Telemetry Engineering Display (TED)
- Persistent Telemetry Storage (PTS)

A server assigned the HOP role accepts LGS WFS or PWFS pixels and performs highly parallelized pixel processing and high-order wavefront reconstruction via a Matrix Vector Multiplication (MVM) to produce DM error vectors.

The server assigned the WCC role accepts OIWFS and ODGW pixels and performs low-order wavefront reconstruction. The WCC also synchronizes and merges the DM error vectors from all of the HOP servers and adds the high-order and low-order information together to compute DM and TTS commands.

The server assigned the TED role is the RTC's interface to the rest of NFIRAOS and other TMT subsystems. It also provides the Engineering GUIs and real-time data displays, as well as the interface to telemetry data. The TED server receives all external commands and dispatches them to the rest of the RTC; this allows the RTC to appear as a single system.

The server assigned the PTS role contains fault tolerant data storage. It receives and stores the telemetry data required by the PSFR. It also stores any other telemetry data that is selected to be saved by the user.

These roles define the functionality of a given server, and therefore the particular hardware requirements for that server, as discussed in Section 3. However, by partitioning the system into these roles, we can somewhat decouple hardware selection from the overall software architecture of the RTC.

### 2.1 AO Mode & Server Roles

The RTC supports three primary modes of operation: LGS MCAO; NGS AO; and Seeing Limited mode. The architecture of the RTC is primarily driven by the requirements of the LGS MCAO mode due to the much greater computational complexity of this mode compared to the others.

In LGS MCAO mode there are up to six HOP servers; one HOP server per LGS WFS. In this mode the WCC will accept all PWFS and other low-order detector pixels for low-order mode reconstruction, and merge the high-order DM error vectors from the six HOP servers, as shown in Figure 2. Additionally, the PWFS acts as a truth wavefront sensor (TWFS) to update the LGS WFS reference vectors.

In NGS AO or Seeing Limited mode there are only four HOP servers; one HOP server per quadrant of the PWFS. In this mode the WCC will accept all low-order detector pixels and merge the high-order DM error vectors from the four HOP servers, as shown in Figure 3.

### 2.2 High-Order Processing (HOP)

The main tasks of a HOP server are: pixel processing; and performing the high-order MVM, as shown in Figure 4. In LGS MCAO a HOP server processes pixels from one LGS WFS, while in NGS AO or Seeing Limited mode HOP server processes pixels for one quadrant of the PWFS. As a result, each HOP server produces a partial high-order DM error vector which must be merged by the WCC.

Figure 5 shows the CPU core assignment on a HOP server for the pixel processing and MVM reconstruction critical threads, along with reading and converting DM shapes into WFS gradients for pseudo-open-loop feedback. In addition, further parallelism is obtained by executing these tasks simultaneously in stages of a pipeline, as pixels arrive.

In order for the MVM to operate on pseudo-open-loop gradients, the DM shape, from previous frames, must be added to the measured gradients. In order for the addition to be possible, the DM shape must first be converted into an equivalent wavefront. This DM shape to wavefront conversion is performed by five threads.

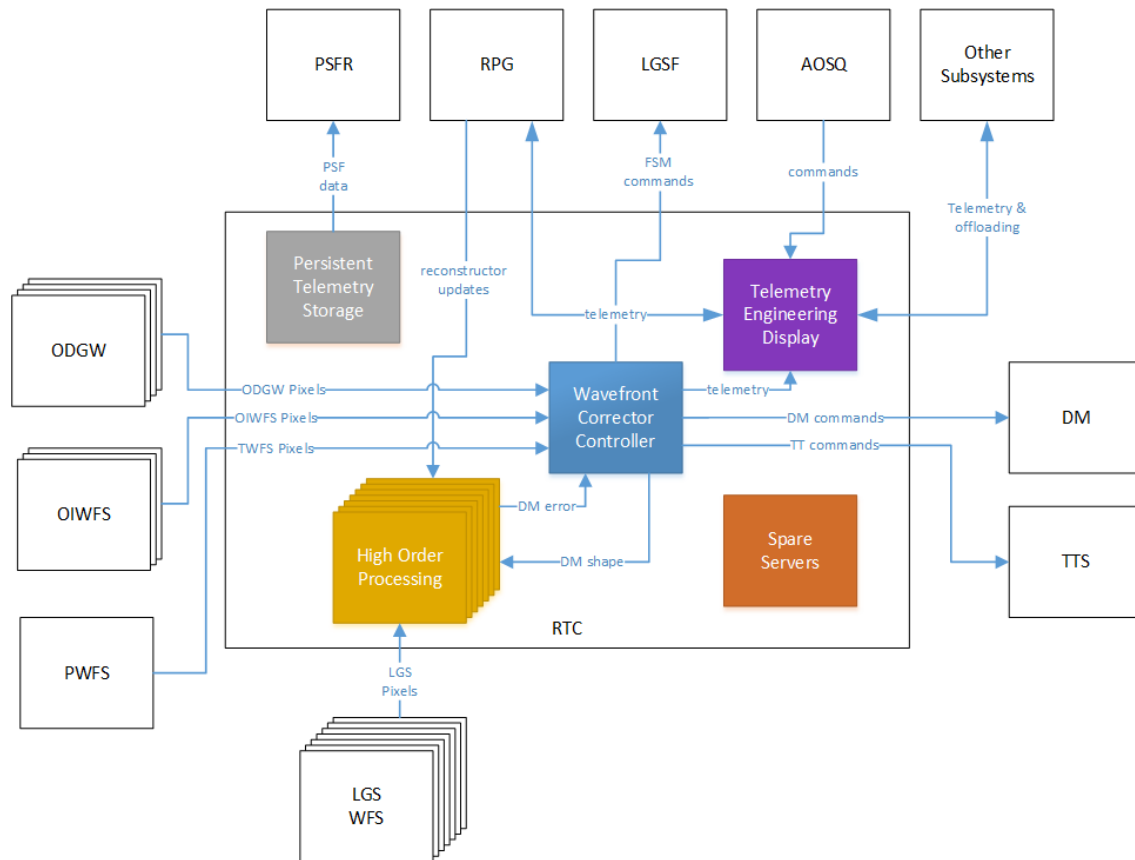


Figure 2. In LGS MCAO mode the six LGS WFSs stream pixels to the six HOP servers while all other low-order and TWFS pixel streams are received by the WCC. Not all data flows are shown in this simplified block diagram.

The first thread reads the DM shape from the WCC server and stores the DM shape in a circular buffer in shared memory. Four threads, one running on each physical CPU, then convert the DM shape into WFS gradients to be added to the measured wavefront error for the next frame.

### 2.3 Wavefront Correction Controller (WCC)

The primary time critical tasks of the WCC server are: performing the final computation of the wavefront correction; and computing the low-order mode correction, as shown in Figure 6. The finalization of the wavefront correction consists of: summing the DM vectors from each of the HOP servers, and adding in the low-order modes; clipping the DM and TTS actuator commands; and finally sending the tip/tilt command to the TTS stage and the DM vectors to the DM electronics. The WCC server also processes pixel data from the low-order detectors: OIWFSs, ODGWs, and the PWFS, and performs low-order mode reconstruction and temporal filtering.

Since the WCC server is responsible for computing the final wavefront correction shape, it also computes the telescope offloading modes for the Telescope Control System. In LGS MCAO the WCC also aggregates the individual tip/tilt (TT) modes for each laser guide star and sends that information to the Laser Guide Star Facility to adjust LGS fast-steering mirrors (FSM).

### 2.4 Telemetry Engineering Display (TED)

The TED is the main interface point between the RTC and several other subsystems. The TED server provides the TMT standard interface to the AOSQ and other TMT services. These interfaces are provided by the RTC Assembly, which handles all commands from the AOSQ and sends subsequent sub-commands to the HOP and

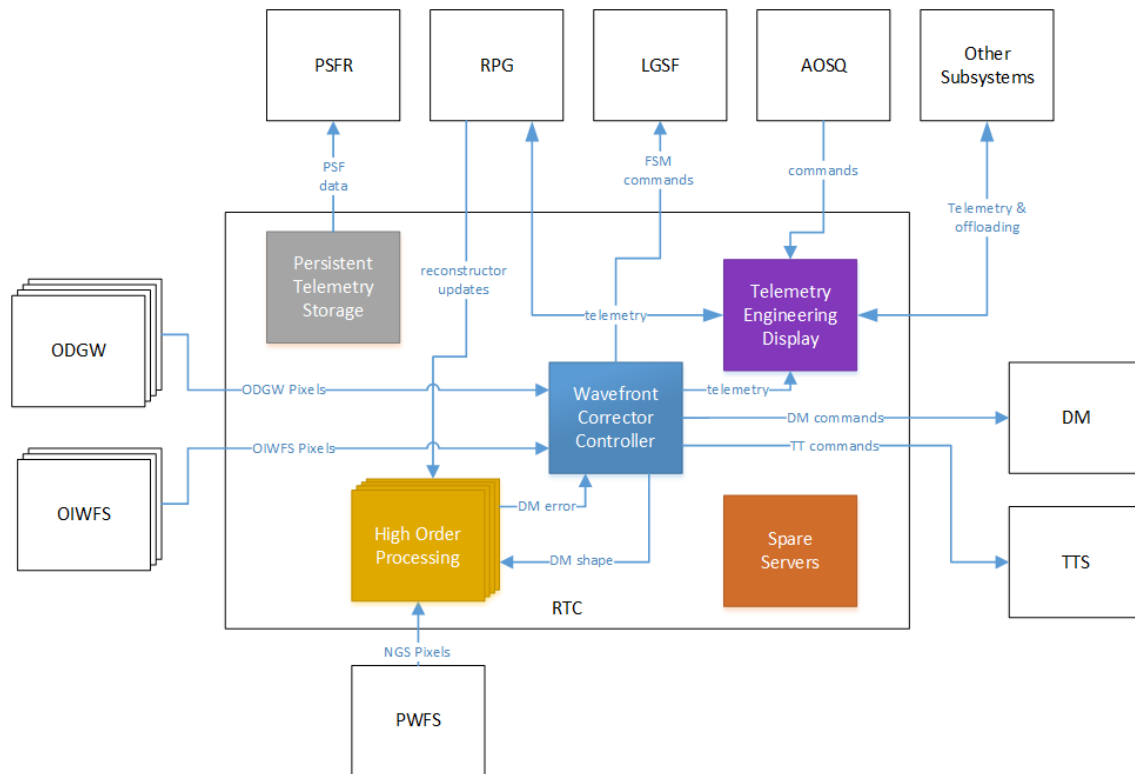


Figure 3. In NGS AO and seeing limited mode the PWFS stream pixels to four HOP servers while all other low-order pixel streams are received by the WCC. Not all data flows are shown in this simplified block diagram.

WCC servers, as shown in Figure 7. The TED server also provides the interface between the rest of the RTC and the following other subsystems:

- RPG - RTC provides telemetry, RPG provides control parameters<sup>†</sup>
- NCC - RTC provides LGS focus and PWFS SSM offloading
- Client Instruments - RTC provides OIWFS POA offloading
- Other Subsystems - RTC provides minor telemetry and offloading
- Engineering GUIs - Engineering control and real-time displays

By acting as an intermediary which aggregates the data, TED also allows the RTC to appear as a single subsystem when multiple RTC servers, like the HOP and WCC, are performing computations.

## 2.5 Persistent Telemetry Storage (PTS)

The role of the Persistent Telemetry Storage (PTS) server is to provide fault tolerant storage for the RTC. The PTS server stores telemetry data required by the PSFR and additional telemetry data saved for future examination.

The fault tolerant storage is achieved by using multiple disks arranged as three distinct RAID-6 arrays, as shown in Figure 8. The PTS also exports two data storage file systems using NFS. These NFS file systems are referred to as: the SSD staging area, which is mounted and written to by RTC servers and read/deleted by the PTS; and the PSFR data storage, which is written to by PTS and mounted/read/deleted by the PSFR.

The SSD staging area is mounted by the RTC servers. The various RTC servers write raw PSFR input data files to the staging area during normal RTC operation. The PTS processes the raw PSFR input files and assembles a FITS file for the PSFR to process. This FITS file is stored on the PSFR data storage partition and

<sup>†</sup>The RPG will send the 160MB high-order reconstruction matrices directly to the HOP servers.

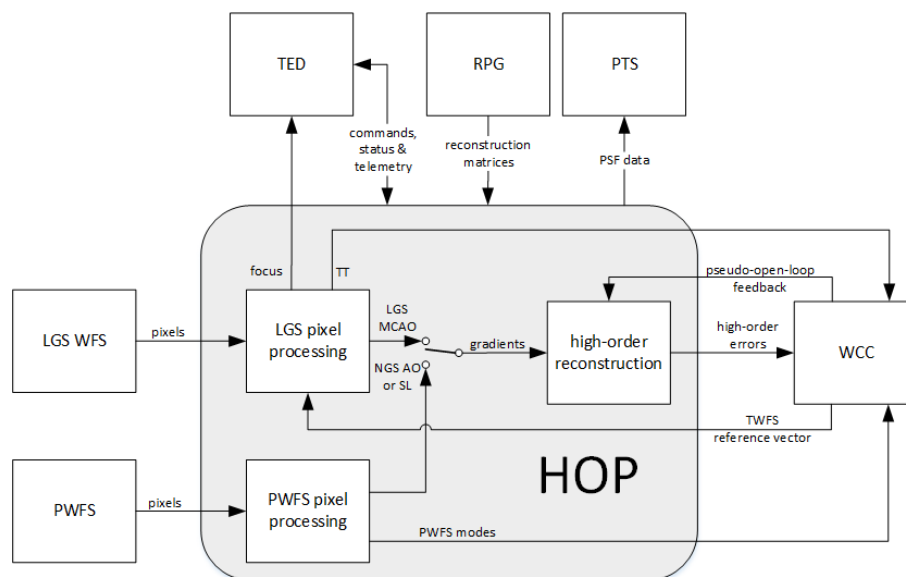


Figure 4. The HOP server role processes high-order pixels and performs high-order reconstruction. In LGS MCAO mode high-order pixels will come from the LGS WFSs, while in NGS AO or Seeing Limited mode the pixels will come from the PWFS. Not all data flows are shown in this simplified block diagram.

is removed by the PSFR once the PSFR processing of the file is completed. When requested by a user, TED instructs the RTC servers to write raw telemetry data to the staging area. The PTS assembles these data into a FITS file, and stores in it the saved telemetry partition on the RAID array. It is expected that the RTC will provide regular data files at a rate of approximately once per second. The rate is chosen as a balance between the number of files and the size of the files.

The PTS server will provide a simple mechanism for downloading the saved telemetry data for examination. This will likely be via an ftp server. The URL to the data file will be provided when the data set to be saved is selected on TED. The PTS server will also run a web server to allow the user to browse the saved telemetry data.

### 3. HARDWARE ARCHITECTURE

All six HOP servers, and additional HOP server spares, will have identical hardware. The WCC and TED, and an additional WCC/TED server spare, will have the same hardware, though different from the HOP servers. The PTS server will also have the same hardware as the WCC and TED servers, but with the addition of a hardware RAID controller and extra storage space.

All of the RTC performance requirements can be met with currently available commercial off-the self hardware.<sup>4</sup> However, new developments in computing and networking should improve performance and reduce latency as discussed below.

#### 3.1 HOP Server Hardware

The most computationally demanding role in the RTC is that of the HOP server. During the RTC Trade Study<sup>5,6</sup> it was demonstrated that one high-end Intel Xeon CPU (E5-2697 V2) was both required and adequate to perform the MVM corresponding to one quadrant of one LGS WFS.

##### 3.1.1 Quad-Socket Server

At the time of the preliminary review the baseline HOP server was as quad-socket machine with E5-4669 V3 CPUs. Each of these E5 CPUs runs at 2.1 GHz and has 18 cores and 45 MiB LLC. However since then we have re-baselined the quad-socket design to use E7-8870 V4<sup>7</sup> CPUs. Each of these E7 CPUs runs at 2.1 GHz and

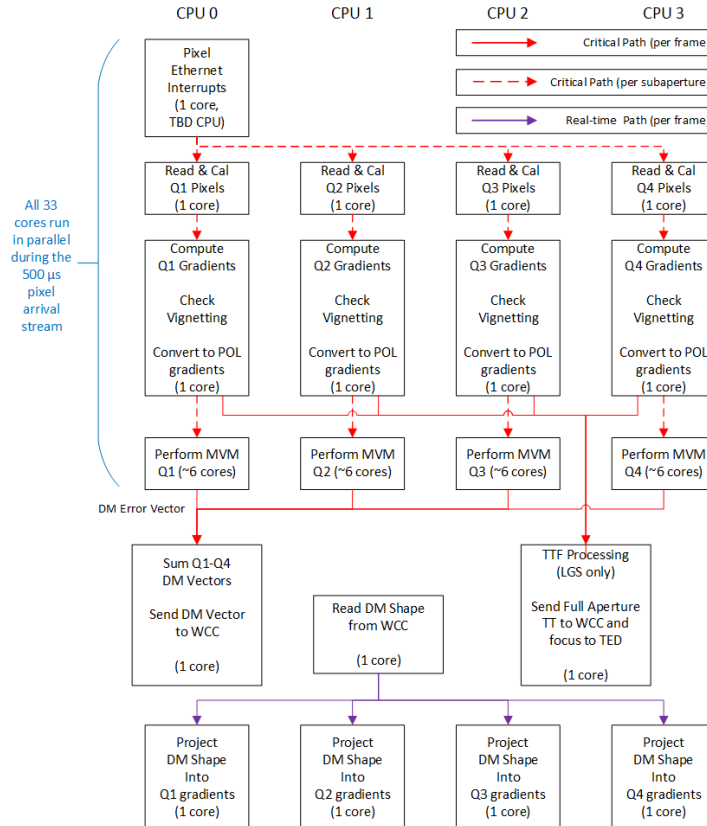


Figure 5. In order to reduce RTC latency, and achieve enough parallelism, the pixel reading and calibration, gradient computation, and the MVM, are all performed in parallel, on a subaperture basis, using one CPU per WFS quadrant. Not all data flows are shown in this simplified block diagram.

has 20 cores and 50 MiB LLC. By moving to the E7 the memory bandwidth increases from 68 GB/s to 102 GB/s, which assists in reading the high-order control matrix from memory. Additionally the E7-8870 V4 is less expensive than the E5-4669 V3, which is a top-end CPU.

The main advantage of the quad-socket design is that one server can process an entire LGS WFS, including both the pixel processing and the MVM. This eliminates some complexity from the software design, and is a natural mapping to the LGS WFS processing. The disadvantage of quad-socket servers is the higher cost for a given amount of performance. In this case, at the time of the preliminary review, the cost is about 50% over that of dual socket machines (excluding disk storage). It is believed that the following benefits outweigh the additional cost of quad socket servers, such as: simplified software design due to having exactly one HOP server per LGS WFS; reduced communication latency; reduced latency for sending LGSF commands; reduced latency for sending DM vectors from HOP servers to WCC; reduced number of machines to administer; reduced number of network connections; and increased reliability due to a smaller number of machines.

### 3.1.2 Knights Landing Xeon Phi

An alternative to quad-socket server solutions, is a design using the Knights Landing Xeon Phi (x200).<sup>8</sup> The Xeon Phi is a bootable (stand-alone) processor that provides massive thread and data parallelism. A Knights Landing Xeon Phi will furnish 72 cores, and will be responsible for processing a single LGS WFS, meaning a single Xeon Phi will act as one HOP server. The Xeon Phi also supports sub-NUMA clustering, such that it can appear as a 4 CPU server, which will minimize code changes between Knights Landing Xeon Phi and quad-socket server implementations.



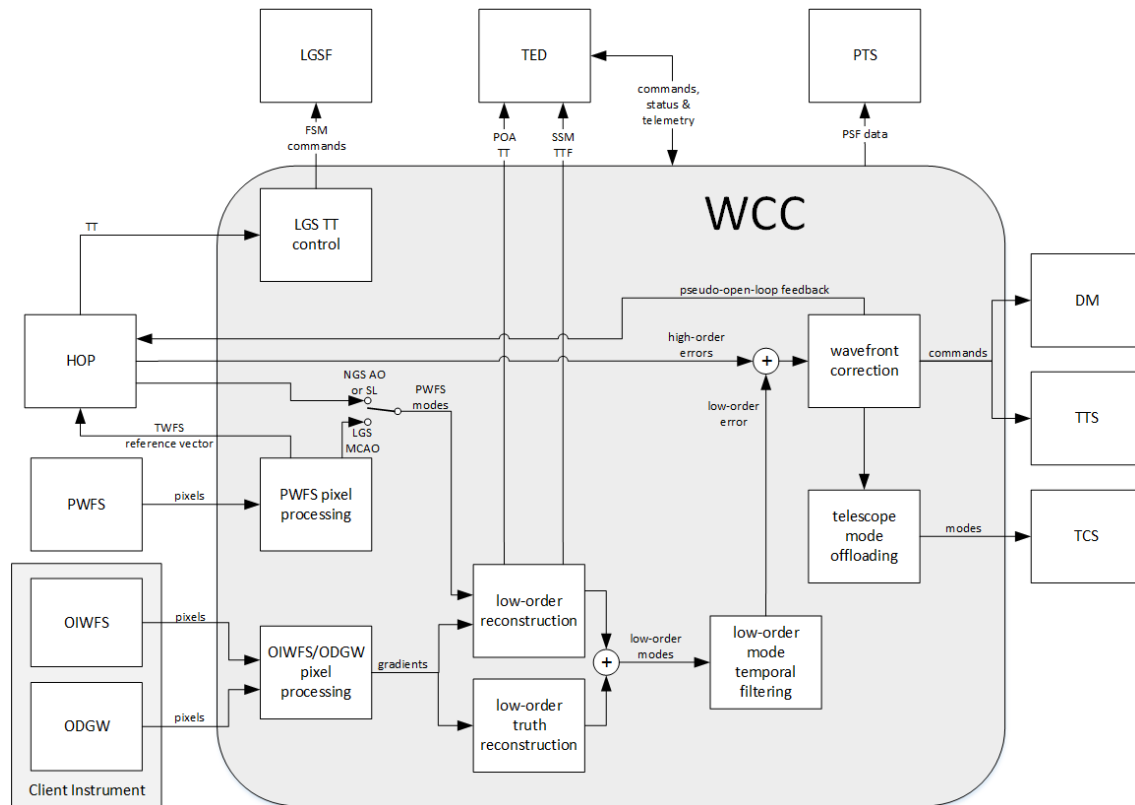


Figure 6. The WCC server role computes the low-order modes and the final wavefront correction. In LGS MCAO mode it also processes PWFS pixels and aggregates LGS TT commands. Not all data flows are shown in this simplified block diagram.

At time of this paper, the Knights Landing Xeon Phi have not been released, therefore benchmarking will be required to verify its suitability as a HOP server. However if benchmarking of the Xeon Phi proves adequate, it will become our preferred hardware selection for the HOP servers.

### 3.2 WCC and TED Server Hardware

At the time of the preliminary review the baseline for the WCC and TED servers was the same hardware as the HOP servers. However, these roles do not require the high core count and cache size needed for the highly parallelized computation of the HOP servers. Therefore a dual-socket server with a modest core count, but higher clock speed of the E5-2643 V4<sup>9</sup> or E5-2667 V4<sup>10</sup> CPUs has been selected as the new baseline for these roles. Each E5-2643 CPU runs at 3.4 GHz, and has 6 cores and 20 MiB LLC, while each E5-2667 CPU runs at 3.2 GHz, and has 8 cores and 25 MiB LLC. The down-selection between these two CPUs will depend on the final thread count, primarily driven by the WCC.

This choice reduces the cost of the WCC and TED servers, however it also means that these servers can no longer share a spare server with the HOP servers and therefore an additional spare for the WCC or TED server is now required.

### 3.3 PTS Server Hardware

The PTS will use the same hardware as the WCC and TED server, as discussed above in Section 3.2. However, the PTS server will also contain a hardware RAID controller and multiple disk drives configured in a fault-tolerant arrangement. The PTS server is also located in the TMT computer room, while all other RTC servers are located in the NFIRAOS electronics enclosure.



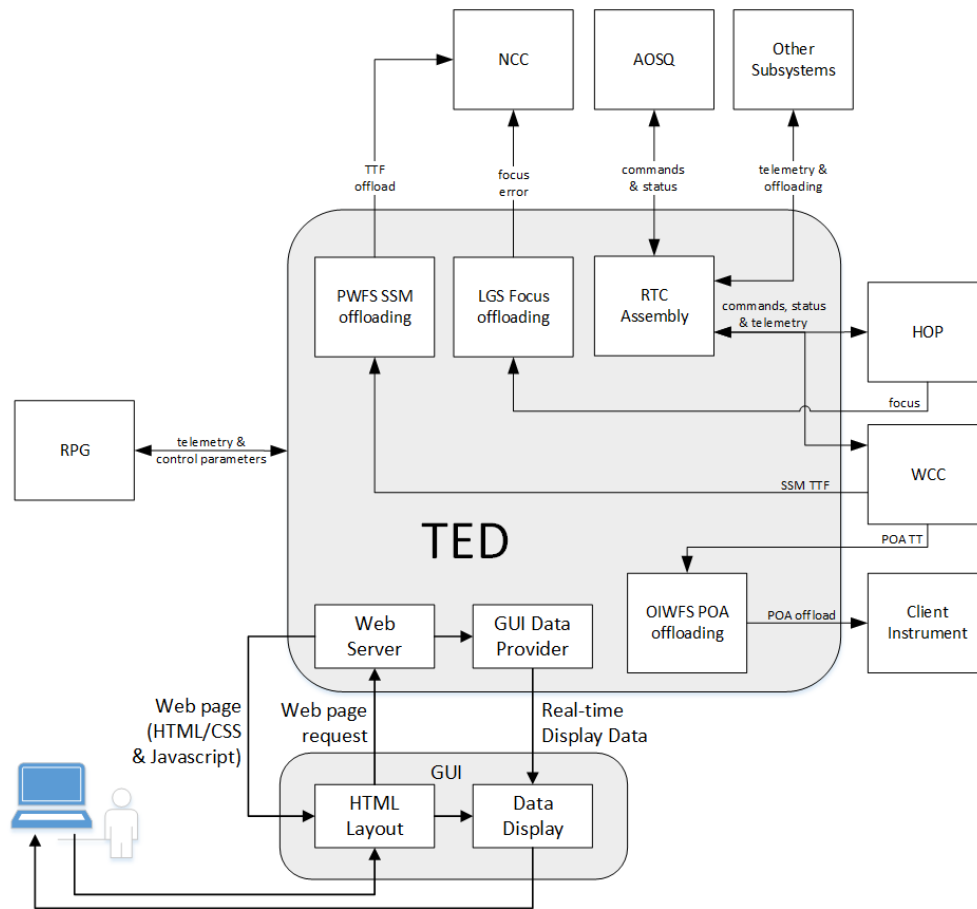


Figure 7. The TED server role manages the interface to the AOSQ and other subsystems and also provides Engineering GUIs and real-time displays. Not all data flows are shown in this simplified block diagram.

### 3.4 RTC Internal Network

At the time of the preliminary review the baseline network design included one 10Gb/40Gb Ethernet switch for both internal and external network traffic. External communication includes: pixels steams; DM and TTS commands; AOSQ commands; status and telemetry; and offloading to other subsystem.

Since the preliminary review the network design has evolved to use Omin Path Fabric for all internal communication.<sup>11</sup> Omni Path provides 100Gb/s and support packet level prioritization for lower latency transport of critical packets. A major source of latency in the RTC design is the transporting of the DM error vectors from the HOP servers to the WCC server, which should be improved by Omni Path connections. Pending benchmarking results, the RTC will use Omni Path fabric for communication between RTC servers, excluding the PTS which will be connected to the TMT infrastructure network. The RTC will still support a 10Gb/40Gb Ethernet switch for all other external communication.

## 4. RTC PIPELINE

The RTC pipeline controls the flow of data throughout the RTC. This pipeline is split into several threads across the RTC servers, yet operates as a single system. The bulk of RTC processing is done within the RTC pipeline, which is principally responsible for pixel processing, wavefront reconstruction, DM and TTS control, and offloading to external systems. The pipeline itself it broken down into server control loops.

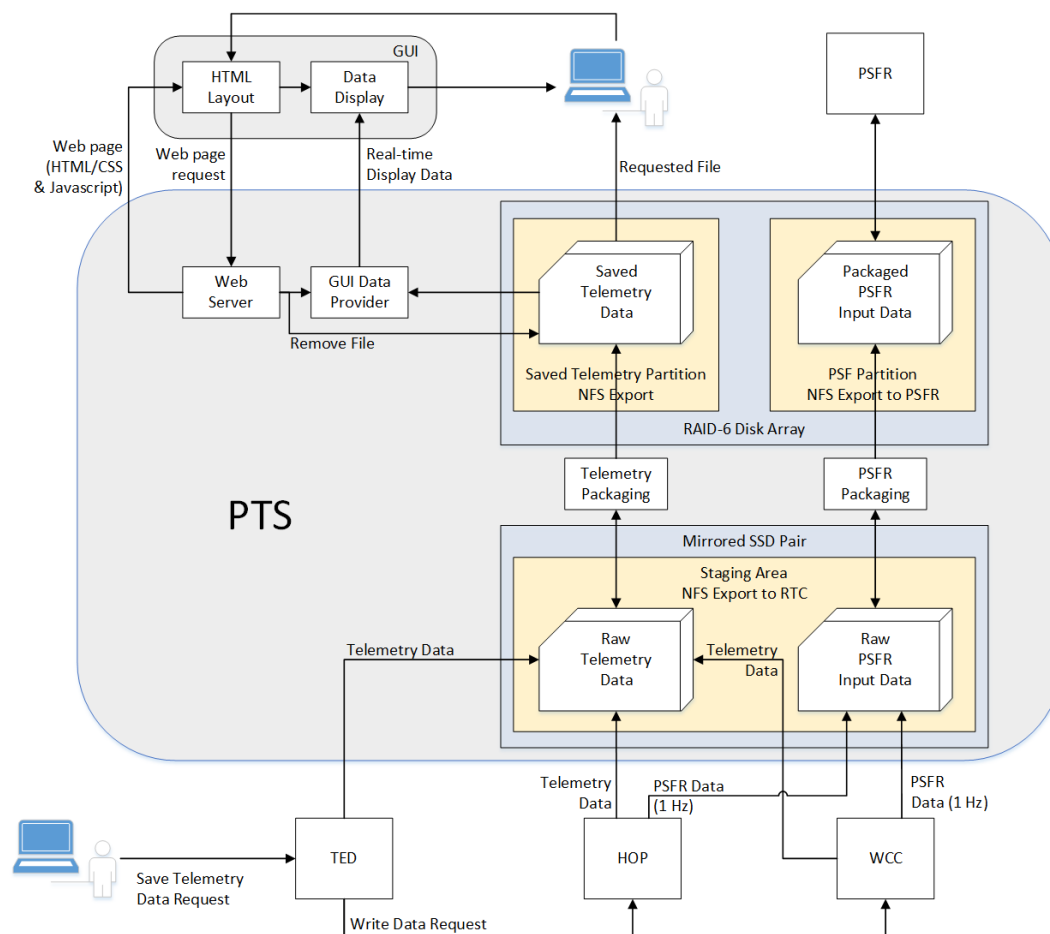


Figure 8. The PTS server role has three distinct RAID-6 arrays. The PTS also exports two NFS data storage. Not all data flows are shown in this simplified block diagram.

#### 4.1 RTC Control Loops

A loop within the context of the RTC refers to an internal operation that will result in the control of or offloading to any device that will affect an input pixel stream. Closing a loop refers to enabling the operation while opening a loop refers to disabling or halting the operation. There are nine major loops with the RTC pipeline. Table 1 shows these loops and the mapping to the RTC server roles responsible for managing the loop.

Regardless of whether the RTC pipeline is active or not, all input pixel streams, either to the HOP servers or the WCC, will be monitored and read, as long as the RTC software is running. These pixels are not processed in any way, and are simply read to avoid build-up in the pixel stream UPD buffers. This provides a convenient quick-look mechanism for raw pixels using the real-time displays, without having to configure the RTC for an observation.

Once the RTC pipeline has been started, pixel and gradient processing is enabled for all active pixel streams. While loops corresponding to a given pixel stream are open, the gradients will be produced by a center of gravity (CoG) method with no optimization. While in this state, low-flux alarms will not be raised, since the guide stars may not even be positioned on the detector yet. The CoG calculation provides a convenient quick-look mechanism for gradients using the real-time displays, without having to close loops.

Table 1. RTC Control Loops

Loop Name	Server Role	Description
high-order	HOP & WCC	high-order control of the DMs and offloads telescope modes to the TCS
low-order	WCC	low-order control of the DMs & TTS, offload telescope modes to the TCS, and send ellipticity and rotator angle errors to the client instrument
LGS focus	HOP & TED	send LGS drift focus offsets to NFIRAOS LGS trombone
LGS TT	HOP & WCC	send LGS TT error commands to the LGSF
LGS dithering	WCC	send dither signal commands to the LGSF and optimize the LGS WFS matched filters
PWFS dithering	WCC	applies a dither signal to the DM or the PWFS FSM and optimize the PWFS optical gain
SSM offload	WCC	offload TTF to the NFIRAOS star selection mechanism (SSM)
POA offload	WCC	offload OIWFS position errors to client instrument OIWFS pick-off arms (POA)
TWFS reference vector update	WCC	update LGS gradient reference vectors based on PWFS measurements

#### 4.1.1 Control Loop Rates

There are several different rates within the RTC pipeline, as a result of the different rates of the input pixel streams, and the desired output rate of wavefront correction commands. In addition to these real-time rates, there are also several background operations that will occur on much longer time scales, and have much looser tolerances on latency and execution timing jitter. These background operations are not discussed in this paper. However, when possible, the background processing will occur during the down-time between frames of the real-time loops.

Once acquisition is complete, and all loops have stabilized, i.e. during steady-state operation, there will be up to four main rates within the RTC control system, as shown in Figure 9.

The wavefront corrector rate is the rate at which DM and TTS commands are sent to their respective drive electronics, and is the principal rate within the RTC, nominally 800 Hz. All other rates divide evenly into the wavefront corrector rate. In LGS MCAO or NGS AO mode the high-order processing rate and the wavefront corrector rate can be the same. In Seeing Limited mode, the low-order and wavefront corrector rates can be the same. However, in any mode, the wavefront corrector rate can be a multiple of the high-order and low-order processing rates, to facilitate low-order mode Kalman filter vibration suppression at the fastest rate possible.

The high-order processing rate is the rate at which the high-order wavefront sensors are streaming pixels to the HOP servers, as well as the rate at which the high-order reconstruction is performed. If the high-order processing rate is slower than the wavefront corrector rate, then the high-order DM error vector stream will be sampled at the faster rate via a circular buffer.

The low-order processing rate is the steady-state rate at which the low-order detectors are streaming pixels to the WCC server, once acquisition is complete, as well as the rate at which the low-order reconstruction is performed. If a Type I or II controller is being used, and the low-order rate is slower than the wavefront corrector rate, then the low-order reconstructed wavefront error stream will need to be sampled at the faster rate. If a Kalman filter is being used, the rate change will be handled implicitly by the Kalman filter. The inputs to the Kalman filter will be applied at the low-order rate while the state updates and filter outputs will be computed at the wavefront corrector rate.

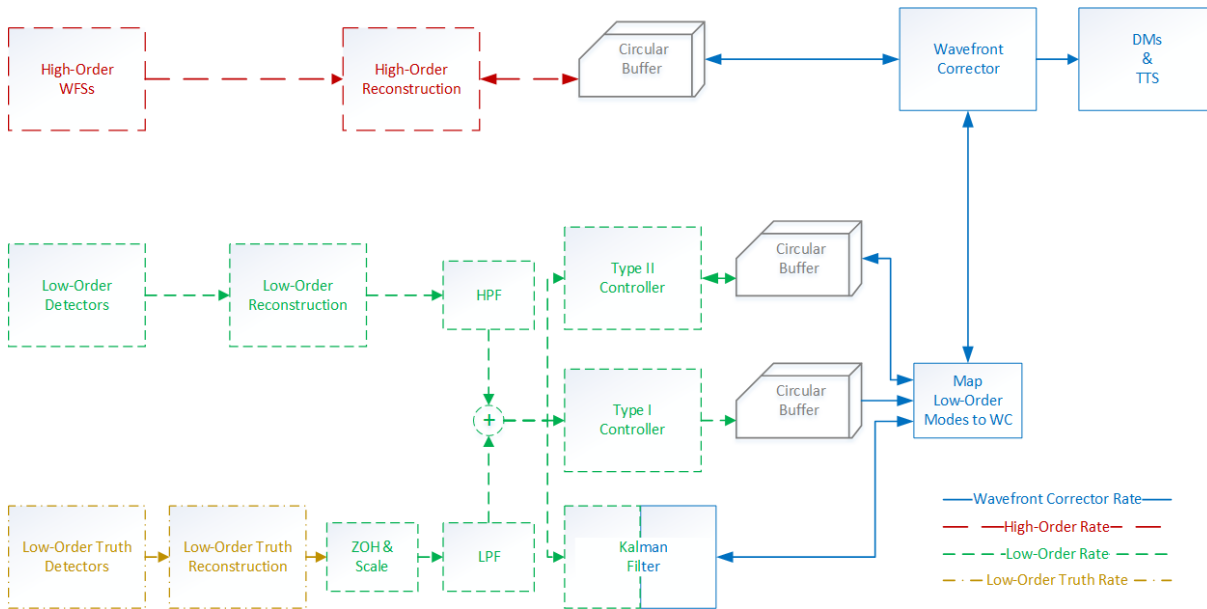


Figure 9. There are four primary loop rate at steady-state. These rates are: the wavefront corrector rate, the high-order processing rate, the low-order processing rate and the low-order truth processing rate. Low-order mode temporal filtering is achieved either by a Type I, Type II controller or a Kalman filter.

The low-order truth rate is the steady-state rate at which the low-order truth detectors are streaming pixel frames to the WCC server, once acquisition is complete. Since the low-order truth rate will always be slower than the low-order rate, a zero-order hold (ZOH) is used to up-sample the truth measurements to the low-order rate, scaled by the ratio between the two rates.

The circular buffer blocks shown in Figure 9 will ensure that the input data are only used once, on the next cycle of the wavefront corrector rate. New high-order and low-order errors are written to the circular buffers, and update an index flag indicating where the latest data resides in the buffer. On each cycle of the wavefront corrector, this index flag will be checked and compared to the value of the flag from the previous cycle. If the value has changed, then there are new data, which the wavefront corrector will apply to the main integrator. If the value has not changed, then there are no new data, and therefore the wavefront corrector will continue with the rest of its computations. These circular buffers operate akin to an impulse function rather than a zero-order hold, such that the full measurement is used as soon as possible, and when no new measurement is available, the output is effectively zero.

#### 4.1.2 Latency and Timing Jitter

A primary driving requirement for the RTC is that the latency, including jitter, between the sending of the pixels from the detector controller and arrival of the the DM commands at the DM electronics shall not exceed 1.2ms, measured over a 10 hour period of operation. The current hardware selection, as discussed in Section 3, can achieve this required performance. However, to provide a robust control system, we have developed a method for mitigating latency and control timing jitter within the RTC pipeline.

To control timing jitter the RTC will define synchronization points in time at which various components of the pipeline can no longer wait for input, either from an external source like the pixels streams, or from a previous component of the pipeline like the DM error vectors from the HOP servers to the WCC. If this time has elapsed, and the required input is not available, then the RTC pipeline must proceed without the data, either by dropping the frame, or operating with reduced performance. These points may also restrict the application of correction from occurring earlier than expected, by delaying the sending of DM commands, to tightly control timing jitter for vibration suppression.

There are three major synchronization points within the RTC, all of which are on the WCC: low-order pixel processing completion; high and low-order wavefront reconstruction completion; and wavefront correction command completion. If low-order pixels do not arrive in a timely manner, the RTC will be forced to drop the frame from the corresponding low-order detector, and use a reduced low-order reconstructor. This is a similar response to losing lock on a low-order guide star, as discussed in Section 4.3.3. If high-order pixels do not arrive in a timely manner, or high-order reconstruction takes longer than expected, the RTC must proceed without any high-order correction for that frame. Finally, if the wavefront correction command computations takes too long, no correction will be applied for that frame.

These synchronization points are defined relative to the arrival of the first pixels, and therefore must account for expected network latency. They will be fully configurable such that, if desired, the RTC will never artificially delay correction and/or extend the deadline past the required latency to avoid dropped frames. At this time we do not see a need to tightly control timing jitter; instead we believe better performance will be achieved by applying the correction as soon as possible, even if it is late.

## 4.2 High-Order Loop Acquisition & Control

Closing the high-order loop will start the high-order control of the DMs. The TTS will also be activated when the high-order loop is started. However, any tip/tilt component will be removed from the net high-order DM error vector by the WCC.

There is no specialized acquisition process for the high-order correction loop; the WCC will simply integrate the net high-order DM error vectors from the HOP servers to drive the DMs. However, if required, the net DM error vector gain may be ramped up over several frames in order to close the high-order loop in a more gradual manner.

If insufficient flux is detected on any high-order WFS, the corresponding HOP server will inform the WCC, and the entire frame will be dropped. However, the high-order loop will remain active. Dropping the frame will effectively freeze the high-order shape on the DM until a complete valid set of high-order WFS frames are received, at which point the main integrator can be updated once again.

## 4.3 Low-Order Loop Acquisition & Control

The low-order loop is responsible for controlling tip, tilt and focus in NGS AO or Seeing mode and tip, tilt, focus and plate scale in LGS MCAO mode. The low-order loop acquisition process is split into up to four tiers. These tiers define the role of a detector within the low-order control path.

- Tier 0: PWFS optionally providing TTF measurements
- Tier 1: required single detector that provides low-order TTF measurements
- Tier 2: 0 to 2 detectors that provide low-order TT or TTF measurements
- Tier 3: 0 to 4 detectors that provide low-order truth TT or TTF measurements

It is anticipated that the tiered low-order loops will be closed in ascending order, waiting for the previous tier to fully acquire before continuing with the next tier. However, this is not a requirement enforced by the RTC. A tier is considered acquired once all detectors within that tier have completed their acquisition process.

Tier 0 is an optional special case in which the PWFS is used for TTF control to facilitate faster acquisition for Tier 1. Tier 1 is the single detector that provides the primary low-order TTF correction. Tier 2 are the up to two additional low-order TT or TTF detectors that, in LGS MCAO mode, provide plate scale correction and additional TTF correction. Tier 3 are the up to four low-order TT or TTF truth detectors, that provide thermal and flexure truth compensation.

The control authority and blending of measurements from the various low-order detectors is achieved by low-order and low-order truth reconstruction matrices, provided by the RPG.

#### 4.3.1 Tier 0 Acquisition

Closing the Tier 0 loop will bypass the normal low-order and low-order truth reconstructors and use TTF measurements from the PWFS to drive the low-order control path. Since Tier 0 always uses the PWFS, there is no RTC related acquisition process, and the TTF modes are passed directly to a static acquisition Type I proportional gain, set by the RPG.

The Tier 0 loop cannot be closed if either Tier 1 or 2 are active. Conversely if Tier 1 or 2 is closed while Tier 0 is active, the RTC will perform a hand-off process to transfer control of the TTF modes from the PWFS to the Tier 1 and/or Tier 2 detectors, and the Tier 0 loop is stopped. If the PWFS is in Tier 1 or 2, then this hand-off is trivial, and closing Tier 1 or 2 will merely consist of starting to apply the PWFS TTF modes to the low-order reconstructor along with the other Tier 1 or 2 gradients.

#### 4.3.2 Tier 1, 2 & 3 Acquisition

The Tier 1 detector provides the measurements for the primary control to low-order TTF modes; either the PWFS or an OIWFS can be assigned to Tier 1. The Tier 2 detectors provide additional TT or TTF measurements, and allow for plate scale mode correction; any of the low-order detectors can be assigned to Tier 2. The Tier 3 detectors provide low-order TT or TTF truth measurements; an OIWFS or ODGW can be assigned to Tier 3.

Though it is expected that the Tier 1 loop will be closed before the Tier 2 loop, this is not a requirement. This possibility facilitates the case where acquisition of the Tier 1 guide star proves to be problematic, potentially because it is actually a galaxy; therefore it may be beneficial to close the Tier 2 loop to get some TT correction to better examine the Tier 1 guide star.

When either the Tier 1 or Tier 2 acquisition process is initiated, including during reacquisition, the low-order path high-pass filter is set to an all-pass filter and the low-order truth path low-pass filter is set to a no-pass filter. In addition, the low-order mode temporal filter will be sent to the static acquisition Type I proportional gain.

The acquisition of Tier 3 guide stars can only be initiated once both Tier 1 and 2 have been fully acquired. If no detectors are assigned to Tier 2, as in NGS AO mode, then only Tier 1 needs to be acquired before acquiring the Tier 3 guide stars. When the Tier 3 acquisition process is initiated, including during reacquisition, the low-order path high-pass filter and low-order truth path low-pass filter are restored to the filter parameters optimized by the RPG.

Once all Tier 1, 2, and 3 guide stars have been fully acquired, the static acquisition Type I proportional gain is replaced with the dynamically updated temporal filters, as specified by the RPG.

#### 4.3.3 Loss of Low-Order Measurement

Loss of measurement in the RTC is defined as any condition that results in a measurement not being available at the time it is required. This could be as a result of: a detector frame gone missing or arriving late; an internal RTC computation thread running slower than required; insufficient flux is detected, i.e. loss of the guide star; or a low-order detector frame is not expected on the current cycle of the low-order rate, either due to slower frame rates during (re)acquisition, or a detector being disabled during non-sidereal repositioning. In any of these cases, the RTC will attempt to continue to operate with the information that is currently available. In the case of low-order mode reconstruction, the RTC will swap in an alternative reconstruction matrix depending on which low-order measurements are available. There are up to three possible low-order detectors (Tiers 1 and 2), resulting in the RTC requiring up to seven variants of the low-order reconstruction matrix. Similarly, there are up to four possible low-order truth detectors (Tier 3), resulting in the RTC requiring up to fifteen variants of the low-order truth reconstruction matrix. Both these matrices are very small, therefore the computational burden of producing all these variants should be trivial. However, in the case of LGS MCAO high-order reconstruction, the RTC is forced to drop the entire high-order frame, in part due to the size of high-order reconstructor.

The overall design philosophy of the RTC has been to provide a robustness control system that delivers the highest-level of correction that is practical given the information that is available. By swapping reconstruction matrices depending on the low-order measurements that are available, the RTC control loops can remain closed, and provide the best possible correction during transient events like reacquisition and non-sidereal repositioning.

This also simplifies the implementation of the initial acquisition process, since the low-order path can be executed at a fixed rate, swapping reconstructors as needed, in response to detector frame rate changes throughout the acquisition process.

#### 4.3.4 Low-Order Loop Rates

During low-order acquisition and reacquisition, various detectors may be operating at different rates. For an OIWFS or ODGW, the rates will continually change depending on the stage of the acquisition process, as discussed in Section 4.4.

During Tier 1 or 2 (re)acquisition, Tier 3 is inactive, which simplifies the (re)acquisition process. The low-order path during Tier 1 or 2 (re)acquisition is shown in Figure 10. In the regular acquisition case, the acquisition processes for both Tier 2 detectors are triggered at the same time after Tier 1 has been fully acquired. However, due to the asynchronous nature of reacquisition, all three detectors within Tier 1 and 2 may be at different stages of the process and therefore running at different rates. Therefore reacquisition is the more general and difficult case. During (re)acquisition the Tier 1 and 2 gradients will be scaled based on a signal gain that corresponds to the current stage of the acquisition process, as discussed in Section 4.4.1. These scaled gradients will be stored in a circular buffer, in which gradients are only used once for low-order reconstruction, on the cycle of the low-order rate that they are initially available. These circular buffers handle the rate change between the current acquisition and the low-order rate, as discussed in Section 4.1.1. By construction, acquisition rates will always be slower or equal to the low-order rate. This results in different low-order reconstructors being swapped in and out throughout the acquisition process, depending on which gradients are available on any give cycle of the low-order loop rate. This swapping of low-order reconstructors is also used to handle missing frames and lost guide stars, as discussed in Section 4.3.3.

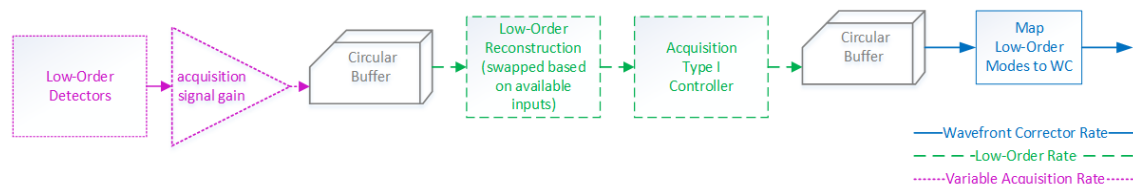


Figure 10. During low-order acquisition each of the, up to three, low-order detectors may operating at different rates. These rate are increase up to the final low-order rate as the acquisition process progresses.

During Tier 3 (re)acquisition, by design, all active Tier 1 and 2 detectors will be fully acquired. In this case, the low-order high-pass filter and low-order truth low-pass filter are set via the parameters optimized by the RPG. By enforcing the restriction that the low-order Tiers (Tier 1 and 2) must be acquired before the low-order truth Tier (Tier 3), the RTC is able to avoid any complication resulting from the temporal splitting of the two paths during (re)acquisition. The complete low-order path during Tier 3 (re)acquisition is shown in Figure 11. Similar to Tier 2, in the regular case, the acquisition process of all Tier 3 detectors is triggered at the same time, however all Tier 3 detectors may be operating at different rates due to asynchronous reacquisition. During (re)acquisition of Tier 3, the gradients will be scaled and zero-order held. This ZOH is used to account for the difference between the acquisition rates and the low-order truth rate. The impulse-like circular buffer scheme discussed in Section 4.1.1 is not used in this case, due to the low-order truth low-pass filter. The two ZOHs in series in the low-order truth path are not an issue since, by construction, Tier 3 acquisition rates will always be slower or equal to the low-order truth rate.

#### 4.3.5 Low-Order Guide Star Hand-off

During non-sidereal tracking, the AOSQ will disable an individual detector within the low-order tiers to facilitate guide star hand-off. This is intended to allow for repositioning of either the OIWFS probe,<sup>12</sup> the PWFS SSM or an ODGW. While a Tier loop is active, the AOSQ will inform the RTC to temporarily ignore one detector stream while the detector or guide window is being repositioned onto the new guide star. This results in a reduction of wavefront information provided to the RTC and a reduced reconstructor will need to be swapped in,



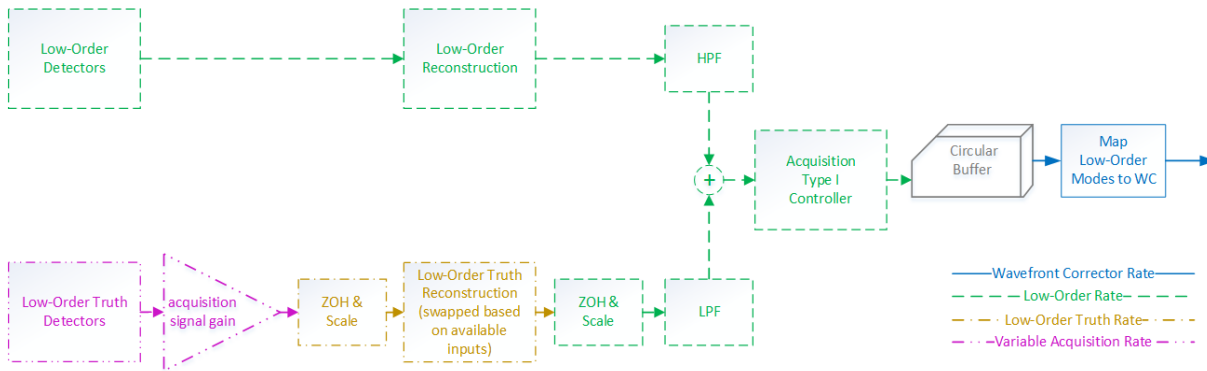


Figure 11. During low-order truth acquisition each of the, up to four, low-order truth detectors may operating at different rates. These rate are increase up to the final low-order truth rate as the acquisition process progresses

as discussed in Section 4.3.3. Once the repositioning is complete, the AOSQ will instruct the RTC to reacquire the new guide star.

## 4.4 OIWFS & ODGW Acquisition & Control

An OIWFS can be configured as either a tip/tilt/focus (TTF) sensor or a tip/tilt (TT) sensor. In TTF mode, an OIWFS can be assigned to Tier 1, 2 or 3, while in TT mode it can only be assigned to Tier 2 or 3. An ODGW can only be used for TT measurements, and therefore can only be assigned to Tier 2 or 3, similar to an OIWFS in TT mode.

### 4.4.1 OIWFS & ODGW Acquisition Tables

While the generalized low-order loop is discussed in Section 4.3, for an OIWFS or ODGW assigned to Tier 1, 2 or 3, a special acquisition process is required. In steady-state an OIWFS or ODGW will send the WCC a 4x4 to 10x10 windowed pixel stream. However, it is expected that a much larger window will be required to initially capture the guide star. At this larger window size, the detectors require additional readout time, and therefore the OIWFS or ODGW pixel streams will be delivered to the WCC much slower than the desired steady-state rate. As a result the guide stars on an OIWFS or ODGW are acquired by having the detector controller step through a pre-defined, observation-specific, acquisition table. The RPG will provide this table to both the detector controller and the RTC. The starting row of an acquisition table will correspond to a large window, and therefore a long readout time, to ensure that the guide star is detectable. As the detector controller steps through the acquisition table, the window size will be reduced, and therefore the readout time will decrease. As a result, the rate at which frames are delivered to the RTC increases as the acquisition process continues. At the final step in the acquisition table, the RTC has locked on the guide star and the corresponding low-order loop is now fully closed.

Prior to the RTCs acquisition of the OIWFS and ODGW guide stars, the AOSQ and the client instrument must first adequately position the OIWFS pick-off arm (POA) and ODGW so that the guide star with be visible on the detector on the first step of the acquisition table.

The structure of the acquisition table, i.e. columns, does not depend on the detector type or tier assignment, and will contain information such as: step identifier; window size; frame rate; exposure time and other detector parameters; pixel processing method; signal gain; and step transitions. However, each detector may require unique acquisition data, i.e. rows, in its table. As an example, if an OIWFS is one of the Tier 2 detectors, and an ODGW is the other, these two detectors, may require different configuration data and even different window sizes. However, both should have the same frame rates at the final step in their respective acquisition tables.

Once the RTC has been instructed to start an OIWFS or ODGW acquisition, it will wait for the detector controller to automatically step through the acquisition table sending the RTC the subsequent pixel frames. The RTC does not require the detector controller to start at the beginning of the acquisition table, as long as

the guide star is detectable in the current frame. The step identifier within the acquisition table, along with other meta data, is sent as a header with each frame of pixels. The RTC will follow along in its copy of the acquisition table, which also specifies a signal gain, to scale the computed gradients, and other control parameters to increase control authority as the acquisition process proceeds.

#### 4.4.2 OIWFS & ODGW Reacquisition

If at any point insufficient flux is detected in an OIWFS or ODGW frame, either during acquisition or after the RTC has acquired lock, the RTC will issue an alarm event to the AOSQ informing it that the guide star has been lost. At this point the AOSQ will either instruct the RTC to open the corresponding loop, or instruct the detector controller to attempt a reacquisition, by jumping to an earlier row in the acquisition table corresponding to a larger window size. If reacquisition is attempted, and the RTC still cannot detect the guide star in the larger window frame, the RTC will issue an additional alarm event informing the AOSQ that the guide star is still lost. While in this state, the RTC will need to swap in a reduced reconstructor to account of the loss on information in the low-order path, as discussed in Section 4.3.3.

This process may continue until the largest window size, corresponding to the starting row in the acquisition table, is reached. If the RTC cannot find the guide star in the largest window, the RTC will issue a final alarm signifying that the guide is still missing, and will likely not be found again, as the detector controller blindly steps through the acquisition table, continually shrinking the window. It is left to the AOSQ to instruct the detector controller and RTC how to proceed at this point. If at any point the guide star is detected within a window, the RTC simply starts applying correction based on the step in the acquisition table, as is done normally during the initial acquisition process.

### 4.5 Acquisition Sequences

From the perspective of the RTC, an acquisition sequence specifies the order in which the RTC loops are closed. The acquisition sequence will depend on the AO mode of the RTC. The command set of the RTC allows for highly configurable acquisition sequences. Only simplified nominal acquisition sequences for LGS MCAO and NGS AO mode are presented here.

#### 4.5.1 LGS MCAO Mode Acquisition Sequence

The baseline acquisition sequence for LGS MCAO mode, from the perspective of the RTC, is split into three major steps. Each of these steps is initiated by the AOSQ.

The first step of the acquisition sequence is to acquire the LGS spots, and close the high-order loop. This is achieved by first starting the RTC pipeline, which starts the basic gradient processing. If insufficient flux is observed on the LGS WFS, the RTC will report a mispointing offload value based on rayleigh backscattering measurements. The AOSQ will guide the LGSF laser launch telescope based on these mispointing values, until sufficient flux is observed on the LGS WFS. At this point the AOSQ will instruct the RTC to close the LGS focus loop, LGS TT loop, and the high-order loop. More detail on the high-order loop is given in Section 4.2. Once these three loops have stabilized (i.e. acquired lock) then the AOSQ may proceed to step two. The AOSQ can enable the LGS dithering if gradient matched filter optimization is desired.

During step two of the sequence, the RTC does not take any direct action, other than maintaining the LGS WFS loops. The AOSQ will instruct NFIRAOS, the client instrument, and the telescope, to position the natural guide stars on the PWFS, OIWFSs and ODGWs. Details of this acquisition process are out of the scope of the RTC design.

Once the natural guide star is positioned on the PWFS, the TWFS reference vector update loop can be started. Additionally, the low-order mode natural guide stars will be acquired in up to 4 parts, acquiring low-order Tiers 0 through 3, as described in Section 4.3. The AOSQ can enable the PWFS dithering if gradient optical gain optimization is desired. The AOSQ can also close SSM and POA offload loops if desired. It may be possible to perform some natural guide star acquisition before acquiring the LGS spots and correcting high-order modes, but that is not the expected sequence as it should be easier to acquire the natural guide stars once the image has already been sharpened by the high-order LGS loops.

#### 4.5.2 NGS AO Mode Acquisition Sequence

The acquisition process for NGS AO mode, from the perspective of the RTC, is split into two major steps. Each of these steps is initiated by the AOSQ.

In the first step the RTC does not take any direct action. The AOSQ will need to instruct NFIRAOS, the client instrument, and the telescope to the position natural guide stars on the PWFS, OIWFSs and ODGWs. Again details of this process are out of the scope of the RTC design.

In the second step, the AOSQ will begin by starting the RTC pipeline. Next, if Tier 0 is specified, the RTC will close the low-order loop performing tip/tilt/focus correction only using the PWFS. Once the PWFS TTF loop, Tier 0, has stabilized (i.e. acquired lock), the AOSQ will instruct the RTC to close the high-order loop. More detail on the high-order loop is given in Section 4.2.

Next, the Tier 1 natural guide stars will be acquired, which will cause the control authority of the TTF modes to be handed off from the Tier 0 PWFS to the Tier 1 detector. Finally, any Tier 3 natural guide stars will be acquired. Details of low-order tier acquisition are given in Section 4.3. The AOSQ can enable the PWFS dithering if gradient optical gain optimization is desired. The AOSQ can also close the SSM and POA offload loops if desired.

### 5. CONCLUSION

The architecture of the RTC has been developed to provide a flexible and robust control system. The role organization allows compartmentalization of components of the RTC pipeline and allows for the streamlined deployment of spare servers in case of hardware failure. Additionally the overall architecture is primarily hardware independent allowing the RTC to take advantage to future hardware development. However the performance requirements can be fully satisfied by currently available commercial off-the self hardware. It is our belief that a CPU-based architecture provides the most flexibility to accommodate future design changes and will be easier to build and maintain, when compared to a FPGA or GPU based design.

The breakdown of the RTC pipeline into smaller control loops provides flexibility in acquisition sequences and provides fine control to the AO Sequencer over the RTC's internal operations. The fault tolerant methodology to control and acquisition ensures that the RTC will continue to operated, though with reduced performance, when frames or guide stars go missing.

### ACKNOWLEDGMENTS

The TMT Project gratefully acknowledges the support of the TMT collaborating institutions. They are the California Institute of Technology, the University of California, the National Astronomical Observatory of Japan, the National Astronomical Observatories of China and their consortium partners, the Department of Science and Technology of India and their supported institutes, and the National Research Council of Canada. This work was supported as well by the Gordon and Betty Moore Foundation, the Canada Foundation for Innovation, the Ontario Ministry of Research and Innovation, the Natural Sciences and Engineering Research Council of Canada, the British Columbia Knowledge Development Fund, the Association of Canadian Universities for Research in Astronomy (ACURA) , the Association of Universities for Research in Astronomy (AURA), the U.S. National Science Foundation, the National Institutes of Natural Sciences of Japan, and the Department of Atomic Energy of India.

### REFERENCES

- [1] Herriot, G., Andersen, D., Atwood, J., Byrnes, P., Caputa, K., Fitzsimmons, J., Hill, A., Lardière, O., Smith, M., and Stocks, J., "NFIRAOS AO for the Thirty Meter Telescope," in [*Adaptive Optics Systems V*], *Proc. SPIE* **9909** (2016).
- [2] Herriot, G., Andersen, D., Atwood, J., Boyer, C., Byrnes, P., Caputa, K., Ellerbroek, B., Gilles, L., Hill, A., Ljusic, Z., Pazder, J., Rosensteiner, M., Smith, M., Spano, P., Szeto, K., Véran, J.-P., Wevers, I., Wang, L., and Wooff, R., "NFIRAOS: first facility AO system for the Thirty Meter Telescope," in [*Adaptive Optics Systems IV*], *Proc. SPIE* **9148**, 914810 (July 2014).

- [3] Boyer, C., Ellerbroek, B., Adkins, S., Andersen, D., Herriot, G., Véran, J.-P., Gilles, L., Wang, L., Otárola, A., Travouillon, T., Wei, K., Tang, J., Bruno, T., Cavaco, J., Pages, H., Siquin, J.-C., Dunn, J., Leach, R., and Schöck, M., “Adaptive optics program update at TMT,” in [*Adaptive Optics Systems V*], *Proc. SPIE* **9909** (2016).
- [4] Smith, M., Kerley, D., Chapin, E., Dunn, J., Herriot, G., Véran, J.-P., Boyer, C., Ellerbroek, B., Gilles, L., and Wang, L., “Thirty Meter Telescope Narrow Field Infrared Adaptive Optics System Real-Time Controller Prototyping Results,” in [*Adaptive Optics Systems V*], *Proc. SPIE* **9909** (2016).
- [5] Smith, M., Kerley, D., Herriot, G., and Véran, J.-P., “Benchmarking hardware architecture candidates for the NFIRAOS real-time controller,” in [*Adaptive Optics Systems IV*], *Proc. SPIE* **9148**, 91484K (July 2014).
- [6] Véran, J.-P., Boyer, C., Ellerbroek, B. L., Gilles, L., Herriot, G., Kerley, D. A., Ljusic, Z., McVeigh, E. A., Prior, R., Smith, M., and Wang, L., “Results of the NFIRAOS RTC trade study,” in [*Adaptive Optics Systems IV*], *Proc. SPIE* **9148**, 91482F (July 2014).
- [7] Intel, “Intel Xeon Processor E7-8870 v4.” [http://ark.intel.com/products/93801/Intel-Xeon-Processor-E7-8870-v4-50M-Cache-2\\_10-GHz](http://ark.intel.com/products/93801/Intel-Xeon-Processor-E7-8870-v4-50M-Cache-2_10-GHz). (Accessed: 9 June 2016).
- [8] Gardner, E., “What public disclosures has Intel made about Knights Landing?.” <https://software.intel.com/en-us/articles/what-disclosures-has-intel-made-about-knights-landing>. (Accessed: 9 June 2016).
- [9] Intel, “Intel Xeon Processor E5-2643 v4.” [http://ark.intel.com/products/92989/Intel-Xeon-Processor-E5-2643-v4-20M-Cache-3\\_40-GHz?q=E5-2643%20v4](http://ark.intel.com/products/92989/Intel-Xeon-Processor-E5-2643-v4-20M-Cache-3_40-GHz?q=E5-2643%20v4). (Accessed: 9 June 2016).
- [10] Intel, “Intel Xeon Processor E5-2667 v4.” [http://ark.intel.com/products/92979/Intel-Xeon-Processor-E5-2667-v4-25M-Cache-3\\_20-GHz?q=E5-2667%20V4](http://ark.intel.com/products/92979/Intel-Xeon-Processor-E5-2667-v4-25M-Cache-3_20-GHz?q=E5-2667%20V4). (Accessed: 9 June 2016).
- [11] Intel, “Intel Omin-Path Fabric 100 Series, The next generation of High Performance Computing (HPC) fabrics.” <http://www.intel.com/content/www/us/en/high-performance-computing-fabrics/omni-path-architecture-fabric-overview.html>. (Accessed: 9 June 2016).
- [12] Chapin, E., Dunn, J., Weiss, J., Gillies, K., Hayano, Y., Johnson, C., Larkin, J., Moore, A., Riddle, R., Sohn, J. M., Smith, R., Suzuki, R., Walth, G., and Wright, S., “The Infrared Imaging Spectrograph (IRIS) for TMT: motion planning with collision avoidance for the on-instrument wavefront sensors,” in [*Software and Cyberinfrastructure for Astronomy IV*], *Proc. SPIE* **9913** (2016).